U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
SYSTEMS DEVELOPMENT OFFICE
TECHNIQUES DEVELOPMENT LABORATORY

TDL OFFICE NOTE 80-4

THE USE OF A MULTIVARIATE LOGIT MODEL FOR PREDICTING QUANTITATIVE PRECIPITATION

Edward A. Zurndorfer and Stephen F. Corfidi

August 1980

THE USE OF A MULTIVARIATE LOGIT MODEL FOR PREDICTING QUANTITATIVE PRECIPITATION

Edward A. Zurndorfer and Stephen F. Corfidi

1. INTRODUCTION

In recent papers, Bocchieri (1979) and Gilhousen (1979) described how the logit model (Brelsford and Jones, 1967) was used to predict the probability of precipitation type (PoPT) and the probability of precipitation (PoP). The results of studies comparing forecasts made from equations developed using the logit model to those made from equations using the Regression Estimation of Event Probabilities (REEP) model (Miller, 1964) indicate that the logit model is indeed useful for predicting precipitation occurrence. Gilhousen (1979) discusses some of the theoretical and empirical reasons for using the logit model for PoP forecasting. In this regard, one of the more important properties of the logit model is that it restricts the probability forecasts between 0 and 1 (as they should be). On the other hand, the REEP forecasts do not have such a restriction. In fact, the REEP derived probability of precipitation amount (PoPA) forecasts (Bermowitz and Zurndorfer, 1979) are occasionally below 0 and on some rare occasions exceed 1. Furthermore, PoPA forecasts less than 0 or greater than 1 suggest that the REEP model suffers from a lack of fit over extreme ranges of predictor values. This is shown, for example, in Figure 1 for a single independent variable. Note how the REEP regression line extends below 0% and exceeds 100% on the probability scale and how well the logit curve estimates probabilities close to 0% and 100%.

In this paper, we describe and present the results of some experiments we performed to compare the logit model and the REEP model for forecasting PoPA and precipitation amount. Forecasts of the probability of $\geq .25$, $\geq .50$, ≥ 1.00, and ≥ 2.00 inches for various projections were made from equations derived from the two models. Furthermore, for both models threshold probabilities that maximize the threat score were derived for each in order to transform the probability forecasts to a "best" categorical forecast. In addition, we verified both the probability and categorical forecasts for the two models.

2. EQUATION DEVELOPMENT

We used the Model Output Statistics (MOS) technique (Glahn and Lowry, 1972) to derive the PoPA equations for both the REEP and logit models. Regionalized equations for the categories $\geq .25$, $\geq .50$, ≥ 1.00, and ≥ 2.00 inches were derived for various projections from the 0000 and 1200 GMT cycles of the Limited-area Fine Mesh (LFM) model (Gerrity, 1977). These projections and the years used to derive the equations are listed in Table 1. Note that equations were derived for both the warm (April-September) and cool (October-March) seasons and that a mixture of LFM-I and LFM-II model data was used in developing the equations. This is because the LFM-II model was implemented in September 1977. Also, for the 6-h projections, we did not derive PoPA equations for the ≥ 2.00 inch category since the event ≥ 2.00 inches was very rare in our developmental sample.

In deriving the REEP PoPA equations, we used a multiple linear screening procedure to select the 12 "best" predictors from a list of about 75 potential predictors. We chose 12 as the number of predictors since experiments have shown that 12 is about the optimum number of terms to use in our PoPA regressions equations (Zurndorfer and Bermowitz, 1976). Equations were derived for nine regions for both the warm and cool seasons.

On the other hand, the version of the logit model which we used does not have a screening option. This model estimates the logit function for up to 10 predictors through the maximum likelihood approach. The initial estimates of the coefficients were obtained through the fitting of a linear discriminant function to the data. The Newton-Raphson procedure of iteration was then used to obtain a final solution of the normal equations.

The first 10 predictors selected by the REEP screening procedure for a given projection, region, and season were included in the appropriate logit equation. Note that for both the REEP and logit equations the same predictors were used for each of the categories $\geq .25$, $\geq .50$, ≥ 1.00, and ≥ 2.00 inches. This is to help ensure consistency among the PoPA forecasts as described in National Weather Service (1980). Also, because the REEP model is linear while the logit model is non-linear, the 10 "best" predictors selected for the REEP equations may not necessarily be the 10 best predictors for the logit equations. However, recent experiments by Gilhousen (1979) indicate little difference in the accuracy of PoP forecasts produced from logit equations with REEP-selected predictors and those produced from logit equations with predictors selected using a logit screening procedure described by Gabriel and Pun (1978).

To transform the PoPA forecasts to categorical forecasts of precipitation amount, we determined threshold probabilities that maximize the threat score. For the REEP model, these probabilities were determined objectively using a procedure described by Bermowitz and Best (1978). Zurndorfer (1980) found that their RC model should be used to determine threshold probabilities for projections less than 24 hours from model run time and that their R model should be used to determine threshold probabilities for projections greater than 24 hours from model run time. Finally, for the logit forecasts, we subjectively determined the threshold probabilities by using a procedure described by Bermowitz and Zurndorfer (1979).

3. RESULTS AND CONCLUSIONS

The results of the PoPA forecast verification are shown in Tables 2 through 6 and those for the categorical forecast verification are shown in Tables 7 through 11. We used the Brier P-score (Brier, 1950) as our verification statistic for the PoPA forecasts. Verification was performed on both dependent and independent data; however, for the categorical forecasts, we verified only on independent data.

We note from Tables 2 through 6 that the REEP PoPA forecasts are consistently better than the logit PoPA forecasts for the $\geq .25$ and $\geq .50$ inch categories for 12-h projections and for the $\geq .25$ inch category for 6-h projections. This was true on both dependent and independent data and for both the warm and cool seasons. On the other hand, for the ≥ 1.00 and ≥ 2.00

inch categories, the logit PoPA forecasts are as good as, and in several cases such as the 0000 GMT 12-24 h projection, better than the REEP PoPA forecasts. This was also true on both the dependent and independent data as well as for both the warm and cool seasons. As we stated in the introduction, the logit model does a better job than the REEP model in estimating probabilities close to 0. Since the event $\geq$ 2.0 inches is rare and PoPA forecasts rarely exceed .10 for this event, we are not surprised to see how well logit did in comparison to REEP for the $\geq$ 2.0 inch category.

Before looking at the results of the categorical forecast verification, we should state one important aspect of this verification. Ideally, when looking at threat scores[1] and biases[2] we would like to see a high threat score (maximum of 1) and a bias equal to, or close to 1. However, due to the difficulty in forecasting precipitation amount, some overforecasting is desirable especially for amounts $\geq$1.00 and $\geq$ 2.00 inches. Therefore, in evaluating the results in Tables 7 through 11 we must look at both the threat score and bias before concluding whether REEP or logit is the better model.

First, for the warm season, we note that the logit biases are very good--close to 1.0 except for the $\geq$ .25 inch category for the 1200 GMT 24-36 h projection and the $\geq$ 2.00 inch category for the 0000 GMT 12-24 h projection. For the 0000 GMT 12-24 h projection results presented in Table 7, we see that the REEP and logit models are about as good with REEP slightly better for the $\geq$ .25 and $\geq$ 2.00 inch categories and logit better for the $\geq$ .50 and $\geq$ 1.00 inch categories. In Table 8, we see for the 1200 GMT 24-36 h projection that the logit biases are all closer to 1 than are the REEP biases while the REEP threat scores are better for the $\geq$ .25 and $\geq$ .50 inch categories and worse for the $\geq$ 1.00 and $\geq$ 2.00 inch categories. Finally, for the warm season, it is clear from Table 9 for the 0000 GMT 6-12 h projection that the logit biases are better than those of REEP for all three categories; however, REEP threat scores are better than the logit threat scores. Overall, we conclude that for the warm season the two models were nearly equal in forecasting precipitation amount.

For the cool season, we see from Tables 10 and 11 a mixture of results. In Table 10, for the 1200 GMT 12-18 h projection, REEP has larger threat scores and biases closer to 1 than logit for the $\geq$ .25 and $\geq$ .50 inch categories; however, logit has the higher threat score and better bias for the $\geq$ 1.00 inch category. Finally, for the 0000 GMT 18-24 h projection (Table 11), we see that logit is somewhat better for the $\geq$ .25 inch category, REEP is better for the $\geq$.50 inch category, and the two models produce the same results for the $\geq$ 1.00 inch category. We conclude that REEP gave somewhat better overall results for forecasting precipitation amount in the cool season.

---

1 Threat score = H/(F + O - H) where H is the number of the correct forecasts and F and O are the number of forecasts and observations of that category, respectively.

2 Bias is the number of forecasts of a category divided by the number of observations of that category. A bias equal to 1.0 means unbiased forecasts.

4. SUMMARY AND OPERATIONAL CONSIDERATIONS

The results of the experiments comparing the REEP and logit models for forecasting PoPA and precipitation amount indicate that the two models are equally good. However, estimating the multivariate logit function is a time consuming procedure from the standpoint of the computer. This is not too surprising given the non-linearity of the logit function. Hence, any procedure (in this case the Newton-Raphson procedure) will require much time for estimating the 40 or so coefficients for the multivariate logit PoPA equations.

Another point to consider is that the logit model requires a time consuming determination of the threshold probabilities for maximizing the threat score. Zurndorfer (1980) indicated that it takes almost 500 minutes on the CPU in order to determine all of the threshold probabilities for an entire season's development. On the other hand, REEP model threshold probabilities can be determined with hardly any additional computer time through an objective procedure which requires only the multiple correlation coefficient (between predictors and predictand) and the climatic frequency of the event.

In summary, while we found the logit model to be useful for forecasting heavy precipitation, we will retain our REEP PoPA system. We do recommend, however, that other developers of forecast equations for rare events such as low ceilings, tornadoes, etc., investigate the use of the logit model.

REFERENCES

Bermowitz, R. J. and D. L. Best, 1978: An objective method for maximizing threat score. Preprints Sixth Conference on Probability and Statistics in the Atmospheric Sciences, Banff, Alberta, Amer. Meteor. Soc., 103-107.

_____________, and E. A. Zurndorfer, 1979: Automated guidance for predicting quantitative precipitation. Mon. Wea. Rev., 107, 122-128.

Bocchieri, J. R., 1979: A new operational system for forecasting precipitation type. Mon. Wea. Rev., 107, 637-649.

Brelsford, W. M. and R. H. Jones, 1967: Estimating probabilities. Mon. Wea. Rev., 95, 570-576.

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. Mon. Wea. Rev., 78, 1-3

Gabriel, K. R. and F. C. Pun, 1978: Binary prediction of weather events with several predictors. Final Report, University of Rochester, Rochester, New York, 99 pp.

Gerrity, J. F., Jr., 1977: The LFM model-1976: A documentation. NOAA Tech. Memo. NWS-NMC-60, NOAA, U.S. Department of Commerce, 68 pp.

Gilhousen, D. G., 1979: Testing the logit model for probability of precipitation forecasting. Preprints Sixth Conference on Probability and Statistics in the Atmospheric Sciences, Banff, Alberta, Amer. Meteor. Soc., 46-48.

Glahn, H. R., and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. J. Appl. Meteor., 11, 1203-1211.

Miller, R. G., 1964: Regression estimation of event probabilities. Tech. Report No. 1, Contract CWB-10704, The Travelers Research Center, Inc., Hartford, Conn., 153 pp.

National Weather Service, 1980: The use of model output statistics for predicting the probability of precipitation amount and precipitation amount categories. NWS Technical Procedures Bulletin No. 283, NOAA, National Weather Service, 12 pp.

Zurndorfer, E. A., 1980: A comparison of objective models for deriving threshold probabilities to maximize the threat score. TDL Office Note 80-3, National Weather Service, NOAA, U.S. Department of Commerce, 5 pp.

______________, and R. J. Bermowitz, 1976: Determination of an optimum number of predictors for probability of precipitation amount forecasting. TDL Office Note 76-17. National Weather Service, NOAA, U.S. Department of Commerce, 7 pp.

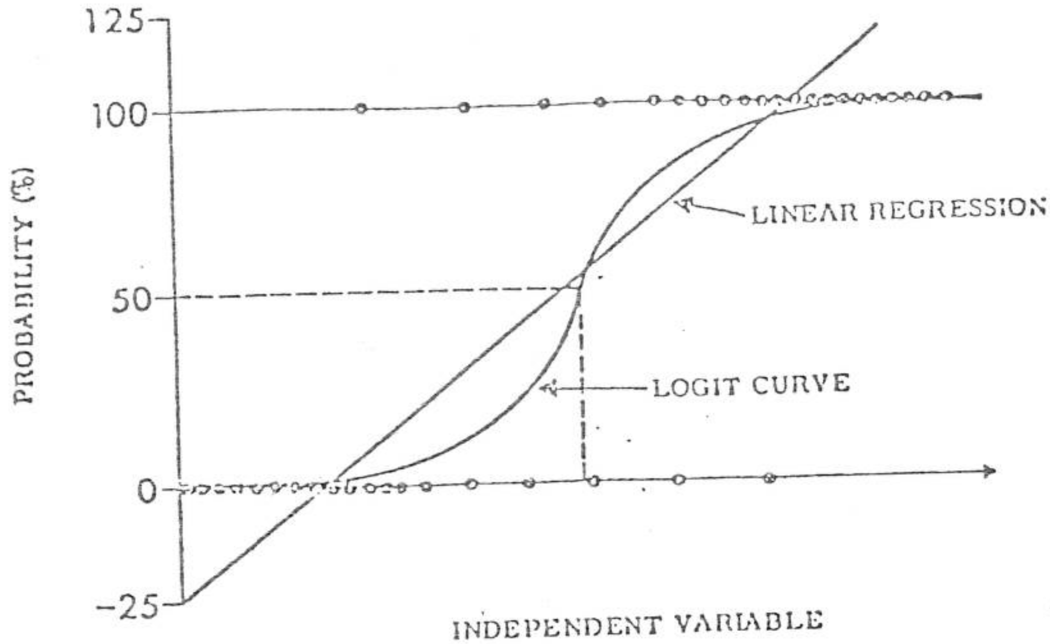

Figure 1. Plot of the logit curve and linear regression for a single independent variable.

Table 1. Projections and number of years of LFM archived forecasts used to derive the PoPA equations.

SEASON	PROJECTION	CYCLE	NUMBER OF YEARS DEVELOPMENTAL DATA
Warm	12-24	00Z	6(1973-1978)
Warm	6-12	00Z	6(1973-1978)
Warm	24-36	12Z	4(1975-1978)
Cool	12-18	12Z	7(1972-73 to 1978-79)
Cool	18-24	00Z	7(1972-73 to 1978-79)

Table 2. Brier Scores for the REEP and logit forecasts. The forecast projection is 12-24 h from 0000 GMT for the warm season. D=Developmental data. I=Independent data (April-September, 1979). Statistics are average scores at 230 cities.

CATEGORY (INCH)	D		I	
	REEP	LOGIT	REEP	LOGIT
≥ .25	.0958	.1079	.1001	.1110
≥ .50	.0548	.0581	.0586	.0613
≥ 1.00	.0219	.0213	.0256	.0245
≥ 2.00	.0055	.0038	.0094	.0066

Table 3. Same as Table 2 except for the projection 6-12 h from 0000 GMT.

CATEGORY (INCH)	D		I	
	REEP	LOGIT	REEP	LOGIT
≥ .25	.0468	.0574	.0537	.0643
≥ .50	.0242	.0254	.0292	.0297
≥1.00	.0077	.0072	.0105	.0098

Table 4. Same as Table 2 except for the projection 24-36 h from 1200 GMT.

CATEGORY (INCH)	D		I	
	REEP	LOGIT	REEP	LOGIT
$\geq$ .25	.0989	.1256	.1045	.1190
$\geq$ .50	.0549	.0643	.0593	.0644
$\geq$1.00	.0214	.0248	.0259	.0257
$\geq$2.00	.0057	.0039	.0097	.0064

Table 5. Same as Table 2 except for the projection 12-18 h from 1200 GMT for the cool season. I=Independent data (October-March, 1978-79).

CATEGORY (INCH)	D		I	
	REEP	LOGIT	REEP	LOGIT
$\geq$ .25	.0377	.0390	.0385	.0400
$\geq$ .50	.0157	.0160	.0164	.0164
$\geq$1.00	.0050	.0039	.0039	.0030

Table 6. Same as Table 5 except for the projection 18-24 h from 0000 GMT.

CATEGORY (INCH)	D		I	
	REEP	LOGIT	REEP	LOGIT
≥ .25	.0389	.0405	.0382	.0395
≥ .50	.0159	.0163	.0157	.0161
≥1.00	.0047	.0036	.0046	.0036

Table 7. Comparative verification of 0000 GMT 12-24 h warm season (April-September) LFM based PoPA categorical forecasts made from the REEP and logit models. Independent data sample consists of the warm season 1979. Statistics are average scores at 230 cities.

CATEGORY (INCH)	THREAT SCORE		BIAS		NUMBER OF CASES
	REEP	LOGIT	REEP	LOGIT	
≥ .25	.248	.245	1.18	1.13	2433
≥ .50	.163	.165	1.21	1.13	1260
≥1.00	.111	.130	1.19	1.01	481
≥2.00	.042	.033	1.42	0.77	124

Table 8. Same as Table 7 except for the 1200 GMT 24-36 h projection.

CATEGORY (INCH)	THREAT SCORE		BIAS		NUMBER OF CASES
	REEP	LOGIT	REEP	LOGIT	
≥ .25	.222	.209	1.65	1.39	2582
≥ .50	.142	.132	1.78	1.21	1322
≥1.00	.087	.092	1.57	1.28	494
≥2.00	.035	.052	2.46	1.09	127

Table 9. Same as Table 7 except for the 0000 GMT 6-12 h projection.

CATEGORY (INCH)	THREAT SCORE		BIAS		NUMBER OF CASES
	REEP	LOGIT	REEP	LOGIT	
≥ .25	.197	.183	1.23	1.11	1171
≥ .50	.139	.128	1.37	1.14	554
≥1.00	.082	.066	1.16	.93	183

Table 10. Same as Table 7 except for the cool season (October-March) 1200 GMT 12-18 h projection. Independent data sample consists of the cool season 1978-79.

CATEGORY (INCH)	THREAT SCORE		BIAS		NUMBER OF CASES
	REEP	LOGIT	REEP	LOGIT	
≥ .25	.258	.250	1.20	1.38	977
≥ .50	.193	.182	1.20	1.42	370
≥1.00	.101	.121	1.45	0.94	62

Table 11. Same as Table 10 except for the 0000 GMT 18-24 h projection.

CATEGORY (INCH)	THREAT SCORE		BIAS		NUMBER OF CASES
	REEP	LOGIT	REEP	LOGIT	
≥ .25	.216	.224	1.45	1.57	950
≥ .50	.149	.139	1.43	2.23	352
≥1.00	.047	.047	.82	.82	73